

УТОЧНЕНИЕ ЛОГАРИФМИЧЕСКОЙ ОЦЕНКИ ТОЧНОСТИ ЖАДНОГО АЛГОРИТМА ДЛЯ ЗАДАЧИ О ПОКРЫТИИ

А.В. Пролубников

к.ф.-м.н., доцент, e-mail: a.v.prolubnikov@mail.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Модифицируя подход, дающий доказательство логарифмической оценки точности жадного алгоритма для задачи о покрытии, мы получаем оценку точности жадного алгоритма для индивидуальной задачи о покрытии. В работе производится сравнение этой оценки с другими оценками такого типа. Даются примеры индивидуальных задач для которых предложенная оценка оказывается существенно лучше, чем оценка, которая может быть получена на основе ЛП-релаксации задачи о покрытии. Используемый подход позволяет разбить множество индивидуальных задач заданной размерности с произвольными весами из \mathbb{R}_+ на классы задач, для которых оценка одинакова, что позволяет показать, что оценка оказывается тем точнее, чем больше минимальная мощность допустимых покрывающих множеств. Показывается также, что с ростом размерности задачи о покрытии доля таких индивидуальных задач растёт.

Ключевые слова: задача о покрытии, жадный алгоритм.

Введение

В задаче о покрытии (ЗПМ), нам дано множество $U = \{1, \dots, m\} \stackrel{\text{dn}}{=} [m]$, и такой набор его подмножеств $S = \{S_1, \dots, S_n\}$, что

$$\bigcup_{i=1}^n S_i = U.$$

Набор $S' = \{S_{i_1}, \dots, S_{i_l}\}$, $S_{i_j} \in S$, называется *покрытием* U , если

$$\bigcup_{j=1}^l S_{i_j} = U.$$

Задана *весовая функция* $w : S \rightarrow \mathbb{R}_+$ ($\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$), $w_i \stackrel{\text{dn}}{=} w(S_i)$ — *вес* множества S_i . Вес набора множеств $S' = \{S_{i_1}, \dots, S_{i_l}\}$ равен сумме весов содержащихся в нём множеств:

$$w(S') = \sum_{j=1}^l w(S_{i_j}).$$

Требуется найти оптимальное покрытие U , то есть необходимо найти покрытие U минимального веса.

Индивидуальная задача о покрытии — это ЗПМ с заданными значениями U , S и w . Пусть \mathcal{A} — алгоритм нахождения приближённого решения для ЗПМ. *Оценкой гарантированной точности* алгоритма \mathcal{A} называется такое значение $\rho_{\mathcal{A}}(m)$, что для любой индивидуальной задачи \mathcal{P} , которая может быть определена на $U = [m]$, мы имеем

$$\frac{w(Cvr)}{w(Opt)} \leq \rho_{\mathcal{A}}(m),$$

где Cvr — это покрытие, полученное с помощью \mathcal{A} для индивидуальной задачи \mathcal{P} , а Opt — оптимальное решение \mathcal{P} .

ЗПМ NP -трудна [1]. В [2] показано, что при условии $P \neq NP$, $\rho_{\mathcal{A}}(m) > (1 - o(1)) \ln m$ для любого приближённого алгоритма \mathcal{A} с полиномиальной вычислительной сложностью. В [3, 4] представлены другие результаты о невозможности получения алгоритмом с полиномиальной вычислительной сложностью приближённого решения с лучшей оценкой гарантированной точности, чем логарифмическая.

Вычислительная сложность жадного алгоритма для ЗПМ составляет $O(m^2n)$. Для этого алгоритма имеет место логарифмическая оценка гарантированной точности [5]:

$$\frac{w(Gr)}{w(Opt)} \leq H(m) \leq \ln m + 1, \tag{1}$$

где Gr — это покрытие, полученное с помощью жадного алгоритма, $H(m) \stackrel{\text{dn}}{=} \sum_{k=1}^m 1/k$. Для индивидуальной ЗПМ мы имеем

$$\frac{w(Gr)}{w(Opt)} \leq H(\tilde{m}), \tag{2}$$

где $\tilde{m} \stackrel{\text{dn}}{=} \max\{|S_i| : S_i \in Opt\}$, и это значение определено для индивидуальной ЗПМ. Поскольку ЗПМ NP -трудна, для того, чтобы получить более точную оценку отношения $w(Gr)/w(Opt)$ для индивидуальной ЗПМ, вместо \tilde{m} мы можем использовать $\bar{m} \stackrel{\text{dn}}{=} \max\{|S_i| : S_i \in S\}$. Для невзвешенного случая ЗПМ, то есть когда $w_i = 1$ для всех S_i , точные верхние и нижние оценки гарантированной точности жадного алгоритма были получены в [6], где показано, что для оценки гарантированной точности $\rho_G(m)$ жадного алгоритма выполняется

$$T_l(m) < \rho_G(m) < T_u(m),$$

где $T_l(m) = \ln m - \ln \ln m - 0.31$, $T_u(m) = \ln m - \ln \ln m + 0.78$.

Рассматривая *индивидуальные задачи* о покрытии множества, которые в дальнейшем, где это подразумевается контекстом, будут именоваться просто *задачами*, мы показываем, как отношение $w(Gr)/w(Opt)$ может быть оценено более точно, нежели предполагается оценкой (1). Представленная нами оценка более точна чем $H(\tilde{m})$ для большинства случаев, для которых значение \tilde{m}

достаточно велико относительно m . Мы приводим такие примеры задач, для которых полученная оценка оказывается более точной, чем оценка, получаемая на основе ЛП-релаксации ЗПМ. Представленная оценка может использоваться для получения нижней границы веса оптимального покрытия для подзадачи задачи о покрытии множества при решении её методом ветвей и границ.

1. Оценивание точности жадного алгоритма для индивидуальной задачи о покрытии множества

Строя покрытие Gr с помощью жадного алгоритма, мы выбираем множества из S в него, исходя из значений $w_i/|S_i|$ весов, назначаемых непокрытым, то есть не входящим в множества из Gr элементам U на момент включения S_i в Gr .

Жадный алгоритм для ЗПМ

Шаг 0. $Gr := \emptyset$.

Шаг 1. Если Gr — покрытие U , то работу алгоритма завершить, иначе перейти на шаг 2.

Шаг 2. Выбрать такое k , что

$$\frac{w_k}{|S_k|} = \min \left\{ \frac{w_i}{|S_i|} : \left(S_i \in S \right) \wedge \left(S_i \not\subseteq \bigcup_{S_j \in Gr} S_j \right) \right\}, \quad (3)$$

$Gr := Gr \cup \{S_k\}$. $S_i := S_i \setminus S_k$ для всех $S_i \notin Gr$. Перейти на шаг 1.

Пусть l — это количество итераций жадного алгоритма, требуемых для покрытия $U = [m]$ в задаче \mathcal{P} . Пусть s_k — это количество элементов U на k -ой итерации, и пусть m_k обозначает количество ещё непокрытых элементов U после выполнения k -ой итерации, $m_0 = m$. Пусть $s = \{s_1, \dots, s_l\}$. Докажем следующую теорему.

Теорема 1.

$$\frac{w(Gr)}{w(Opt)} \leq G(s) = H(m) - \Delta(\mathcal{P}), \quad (4)$$

где

$$G(s) \stackrel{\text{dn}}{=} \sum_{k=1}^l \frac{s_k}{m_{k-1}}.$$

$\Delta(\mathcal{P}) \geq 0$, то есть $G(s)$ уточняет оценку $H(m)$, и $\Delta(\mathcal{P}) = 0$ только в том случае, когда требуется m итераций жадного алгоритма для получения покрытия U .

В случае наличия нескольких множеств из S с равным значением отношения (3) выбор множеств в Gr на итерации жадного алгоритма будем производить лексикографически. Вместо обозначения $G(s)$ будем использовать также обозначение $G(\mathcal{P})$, поскольку задача \mathcal{P} единственным образом определяет последовательность s .

Оценка (4) является уточнением оценки (1). Мы доказываем (4), модифицируя хорошо известное доказательство (1), представленное, например, в [7]. Но оценивая отношение $w(Gr)/w(Opt)$, вместо мажорирования весов, которые множества из Gr назначают элементам U , мы мажорируем веса самих множеств. В результате мы получаем оценку (4), которая оказывается более точной чем $H(\bar{m})$ для большой доли индивидуальных ЗПМ при заданном m .

Предположим, что перед выполнением k -ой итерации жадного алгоритма непокрытым остаётся подмножество U_k элементов множества U , $m_k = |U_k|$. После выполнения шага 2 ($k-1$)-ой итерации все множества из S содержат только элементы U_k . Предположим, что жадный алгоритм выбирает в соответствии с (3) множество S_k на k -ой итерации. Тогда, как мы покажем далее,

$$\frac{w(S_k)}{|S_k|} \leq \frac{w(Opt_k)}{m_k},$$

где Opt_k — это оптимальное покрытие U_k , которое может быть получено из модифицированных множеств из S . Мы используем это неравенство для того, чтобы доказать (4). Оно также используется для доказательства (1) и (2), но мы используем его для доказательства оценки точности другого типа. А именно, мы не оцениваем гарантированную точность жадного алгоритма, находя такую оценку для худшего случая индивидуальной задачи заданной размерности, а оцениваем точность жадного алгоритма при решении с его помощью данной индивидуальной задачи.

Пусть набор множеств $\{S_1, S_2, \dots, S_l\}$ — покрытие Gr для ЗПМ \mathcal{P} с $U = [m]$. Пусть $Opt = \{A_1, A_2, \dots, A_r\}$, $A_i \in S$ — оптимальное покрытие и пусть $a_i \stackrel{\text{dn}}{=} w(A_i)$.

Лемма 1.

$$\frac{w(S_1)}{|S_1|} \leq \frac{w(Opt)}{m}.$$

Доказательство. Пусть

$$A'_1 = A_1, A'_2 = A_2 \setminus A_1, \dots, A'_r = A_r \setminus \bigcup_{j=1}^{r-1} A_j.$$

$A'_j \cap A'_k = \emptyset$ для $j \neq k$. Перенумеруем множества A'_j в соответствии с убыванием значения отношения $a_j/|A'_j|$:

$$\frac{a_1}{|A'_1|} \leq \frac{a_2}{|A'_2|} \leq \dots \leq \frac{a_r}{|A'_r|}.$$

Поскольку для положительных a, b, c, d выполняется

$$\frac{a}{c} \leq \frac{b}{d} \Rightarrow \frac{a}{c} \leq \frac{a+b}{c+d},$$

то

$$\frac{a_1}{|A'_1|} \leq \frac{\sum_{j=1}^r a_j}{\sum_{j=1}^r |A'_j|} = \frac{w(Opt)}{m}.$$

Учитывая (3), получаем

$$\frac{w(S_1)}{|S_1|} \leq \frac{a_1}{|A_1|} = \frac{a_1}{|A'_1|} \leq \frac{w(Opt)}{m}.$$

■

Докажем Теорему 1.

Доказательство. Рассмотрим выбор множества в Gr на k -ой итерации жадного алгоритма как выбор первого множества в покрытие той части множества U , которая ещё не покрыта перед выполнением k -ой итерации. Это эквивалентно рассмотрению задачи с модифицированными множествами из S . Веса этих множеств те же.

Пусть на k -й итерации жадного алгоритма выбрано множество \widehat{S}_k , где

$$\widehat{S}_k = S_k \setminus \bigcup_{j=1}^{k-1} S_j.$$

Поскольку $s_k = |\widehat{S}_k|$, по доказанной лемме мы имеем:

$$w_1 \leq \frac{s_1}{m_0} w(Opt), \quad w_2 \leq \frac{s_2}{m_1} w(Opt_2), \quad \dots, \quad w_l \leq \frac{s_l}{m_{l-1}} w(Opt_l).$$

Так как $w(Opt_l) \leq \dots \leq w(Opt_2) \leq w(Opt)$, то

$$\begin{aligned} w(Gr) &= \sum_{k=1}^l w_k \leq \frac{s_1}{m_0} w(Opt) + \frac{s_2}{m_1} w(Opt_2) + \dots + \frac{s_l}{m_{l-1}} w(Opt_l) \leq \\ &\leq \left(\frac{s_1}{m_0} + \frac{s_2}{m_1} + \dots + \frac{s_l}{m_{l-1}} \right) w(Opt) = \\ &= \left(H(m) - H(m) + \frac{s_1}{m_0} + \frac{s_2}{m_1} + \dots + \frac{s_l}{m_{l-1}} \right) w(Opt) = (H(m) - \Delta(\mathcal{P}))w(Opt), \end{aligned}$$

где

$$\Delta(\mathcal{P}) = H(m) - \left(\frac{s_1}{m_0} + \frac{s_2}{m_1} + \dots + \frac{s_l}{m_{l-1}} \right).$$

Полагая $m_l = 0$ и производя суммирование по i в убывающем порядке, мы получим

$$\Delta(\mathcal{P}) = \sum_{k=1}^l \left(\sum_{i=m_{k-1}}^{m_k+1} \frac{1}{i} - \frac{s_k}{m_{k-1}} \right).$$

Поскольку

$$\begin{aligned} \Delta(\mathcal{P}) &= \sum_{k=1}^l \left(\sum_{i=m_{k-1}}^{m_k+1} \frac{1}{i} - \frac{s_k}{m_{k-1}} \right) = \sum_{k=1}^l \left(\sum_{i=m_{k-1}}^{m_k+1} \frac{1}{i} - \sum_{i=m_{k-1}}^{m_k+1} \frac{1}{m_{k-1}} \right) = \\ &= \sum_{k=1}^l \sum_{i=m_{k-1}}^{m_k+1} \left(\frac{1}{i} - \frac{1}{m_{k-1}} \right), \end{aligned}$$

мы имеем $\Delta(\mathcal{P}) \geq 0$ и $\Delta(\mathcal{P}) = 0$ только в том случае, если $m_{k-1} = m_k + 1$, то есть когда $s_k = 1$ для всех $k = \overline{1, l}$. ■

2. Оценивание точности жадного алгоритма с помощью $G(s)$ и других оценок

Пусть \mathcal{P} — ЗПМ с $U = [m]$, $S \subseteq 2^U$. По доказанной теореме мы можем для \mathcal{P} оценить отношение $w(Gr)/w(Opt)$ используя $G(s)$, где s — последовательность, получаемая в ходе работы жадного алгоритма на \mathcal{P} . Пусть $m(s) \stackrel{\text{dn}}{=} \max\{|s_i| : s_i \in S\}$.

2.1. Сравнение $G(s)$ с $H(\bar{m})$

Сравним $G(s)$ с $H(\bar{m})$ (не с $H(m)$), рассматривая их как два различных способа получения оценки на $w(Gr)/w(Opt)$ для индивидуальной ЗПМ. Мы сравниваем $G(s)$ с $H(\bar{m})$ для всех ЗПМ, которые могут быть определены на $U = [m]$ для $m = 10, 35$, используя разбиение задач на классы \mathcal{C}_s , которые могут быть заданы на U .

Классы задач \mathcal{C}_s . Для каждой последовательности $s = \{s_1, \dots, s_l\}$ имеется класс таких ЗПМ \mathcal{C}_s , что жадный алгоритм в ходе своей работы на них даёт s . Например, жадный алгоритм даёт последовательность s для ЗПМ следующего вида. Пусть $l > 1$, $q_i = \sum_{j=1}^i s_j$, $i = \overline{2, l}$, $q_1 = 0$. Пусть $S = \{S_1, \dots, S_l, A\}$, где $S_i = \{q_i + 1, \dots, q_i + s_i\}$, $A = \{1, 2, \dots, m\}$. Пусть $w(S_i) = |S_i| = s_i$ для $i = \overline{1, l-1}$, $w(S_l) = |S_l| + 1$, $w(A) = m + \varepsilon$, $0 < \varepsilon < 1$. Применяя жадный алгоритм, для такой задачи мы получаем покрытие, которое состоит из множеств S_1, \dots, S_l . Вес этого покрытия равен $m + 1$, тогда как вес оптимального покрытия (вес множества A) равен $m + \varepsilon$. Таким образом, для ЗПМ указанного вида в ходе работы жадного алгоритма будет получена последовательность s .

Чтобы сравнить $G(s)$ и $H(\bar{m})$, для каждого m мы разбиваем множество всех индивидуальных ЗПМ, которые могут быть заданы на $U = [m]$ на классы

\mathcal{C}_s . Поскольку $m(s) \leq \bar{m}$ для любой $\mathcal{P} \in \mathcal{C}_s$, значение $H(m(s))$ не превышает значения $H(\bar{m})$. Это значит, что подсчитывая количество последовательностей s , для которых $G(s) < H(m(s))$, мы получаем оценку на количество таких последовательностей s , для которых выполняется неравенство $G(s) < H(\bar{m})$. Таким образом для данного значения m мы оцениваем долю таких классов \mathcal{C}_s , что $G(s) < H(\bar{m})$ для $\mathcal{P} \in \mathcal{C}_s$.

Доля классов \mathcal{C}_s , для которых $G(s) < H(\bar{m})$. Пусть $\mu(s) \stackrel{\text{dn}}{=} m(s)/m$. В табл. 1 в соответствии с $\mu(s)$, принадлежащим заданным интервалам, представлены доли от количества всех классов \mathcal{C}_s , для которых оценка $G(s)$ оказывается более точной, чем оценка $H(m(s))$. Расчёты, проведённые для всех возможных случаев s для $m = \overline{10, 35}$, показывают, что доля таких классов растёт с ростом $\mu(s)$. Также они показывают, что эта доля растёт с ростом m .

Уточнение оценки $H(\bar{m})$ оценкой $G(s)$ при $G(s) \leq H(\bar{m})$. Данные, представленные в табл. 2 и 3, показывают, насколько оценка $G(s)$ может быть более точной, чем $H(m(s))$ для различных значений $\mu(s)$. Для таких s что $G(s) < H(m(s))$, пусть

$$\Delta(s) \stackrel{\text{dn}}{=} (H(m(s)) - G(s)) / H(m(s)) \times 100,$$

то есть $\Delta(s) = H(m(s)) - G(s)$ — величина уточнения $H(m(s))$, получаемого с помощью оценки $G(s)$ в процентах от $H(m(s))$. В табл. 2 и 3 представлены средние и максимальные значения этой величины для $m = \overline{10, 35}$ и различных значений $\mu(s)$: тогда как $H(m(s)) \in [2,93; 4,15]$ для определённых на $[m]$ последовательностей s , при достаточно больших значениях $\mu(s)$ разница $\Delta(s)$ может достигать десятков процентов.

Подводя итог, мы можем сказать, что $H(\bar{m})$ оказывается более точной, чем $G(s)$, когда \bar{m} достаточно мало, то есть когда S содержит только множества небольшой мощности. Однако если в S присутствуют множества достаточно большой мощности, которые включаются в Gr жадным алгоритмом, то оценка $G(s)$ оказывается более точной, чем $H(\bar{m})$. Для случая $\bar{m} = m$ и $s_1 = m$, $G(s) = 1$ и $H(\bar{m}) = H(m)$, то есть оценка $G(s)$ отношения $w(Gr)/w(Opt)$ точна, тогда как значение $H(\bar{m})$ достигает своего максимального значения.

2.2. $G(s)$ и оценки на основе ЛП-релаксации ЗПМ

ЗПМ может быть поставлена как задача целочисленного линейного программирования. Решение таким образом поставленной ЗПМ представляет собой булев вектор размерности n , для которого равенство его компоненты единице эквивалентно включению соответствующего множества в покрытие, равенство нулю — невключению. *Дробное покрытие* — это допустимое решение задачи линейного программирования, полученной *ЛП-релаксацией* целочисленной индивидуальной ЗПМ, то есть отбрасыванием условия целочисленности

Таблица 1. Доля классов C_s , для которых $G(s) < H(m(s))$

m	$\mu \in (0; 0,2]$	$\mu \in (0,2; 0,4]$	$\mu \in (0,4; 0,6]$	$\mu \in (0,6; 0,8]$	$\mu \in (0,8; 1]$
10	0	13,5	64,8	100	100
11	0	12,4	54,7	92,5	100
12	0	9,9	52,9	97,5	100
13	0	13,9	64,5	93,3	100
14	0	9,1	56,1	96,2	100
15	0	14,5	69,6	99,0	100
16	0	12,9	62,6	96,2	100
17	0	11,6	58,7	98,3	100
18	0	12,4	65,0	95,7	100
19	0	11,0	61,2	97,7	100
20	0	19,5	67,4	99,1	100
21	0	15,5	63,2	97,2	100
22	0	11,8	59,6	98,7	100
23	0	11,1	64,7	96,7	100
24	0	10,1	61,3	98,1	100
25	0	18,6	68,5	99,2	100
26	0	16,9	64,2	97,8	100
27	0	15,1	59,7	98,8	100
28	0	14,1	65,9	97,4	100
29	0	12,0	62,8	98,5	100
30	0,5	17,2	67,2	99,2	100
31	0,2	15,8	64,0	98,1	100
32	0,1	14,5	61,0	98,9	100
33	0	13,5	65,0	97,8	100
34	0	12,5	62,2	98,6	100
35	0,9	15,8	66,9	99,3	100

компонент вектора решений. Имея оптимальное дробное покрытие $Opt_{LP} \in \mathbb{R}^n$ для задачи \mathcal{P} , так как $w(Opt_{LP}) \leq w(Opt)$, мы можем получить оценку сверху на $w(Gr)/w(Opt)$:

$$\frac{w(Gr)}{w(Opt)} \leq \frac{w(Gr)}{w(Opt_{LP})}. \tag{5}$$

Таблица 2. Уточнение оценки $H(m(s))$ оценкой $G(s)$

m	$\mu \in (0; 0,2]$		$\mu \in (0,2; 0,4]$		$\mu \in (0,4; 0,6]$	
	ср. зн. Δ	max Δ	ср. зн. Δ	max Δ	ср. зн. Δ	max Δ
10	0	0	12,3	18,4	21,3	42,9
11	0	0	8,9	14,9	20,1	40,6
12	0	0	6,7	12,0	19,1	45,4
13	0	0	10,9	24,2	19,1	43,6
14	0	0	10,2	21,8	20,1	47,4
15	0	0	13,9	30,6	20,7	50,5
16	0	0	13,1	28,6	19,7	49,2
17	0	0	11,7	26,8	19,1	51,8
18	0	0	12,8	33,6	20,0	50,7
19	0	0	12,4	32,0	19,5	52,9
20	0	0	12,7	37,5	20,8	54,9
21	0	0	12,2	36,0	19,8	54,0
22	0	0	12,3	34,8	19,2	55,7
23	0	0	13,5	39,3	19,9	54,9
24	0	0	13,2	38,1	19,4	56,4
25	0	0	12,9	42,0	19,8	57,8
26	0	0	12,3	40,9	19,2	57,1
27	0	0	12,0	39,9	19,0	58,4
28	0	0	12,3	43,2	19,2	57,7
29	0	0	12,2	42,3	18,7	58,9
30	2,7	6,8	13,1	45,2	19,4	59,9
31	2,2	5,5	12,7	44,4	18,8	59,4
32	1,7	4,3	12,4	43,6	18,4	60,4
33	1,3	3,1	12,4	46,2	18,9	59,9
34	0,9	2,0	12,2	45,4	18,5	60,8
35	5,5	11,9	13,3	47,7	18,8	61,6

Пусть $R(\mathcal{P}) \stackrel{\text{dn}}{=} w(Gr)/w(Opt_{LP})$.

Разрыв целочисленности $IG(\mathcal{P})$ — это значение, характеризующее точность приближённого решения, получаемого решением ЛП-релаксации цело-

Таблица 3. Уточнение оценки $H(m(s))$ оценкой $G(s)$

m	$\mu \in (0,6; 0,8]$		$\mu \in (0,8; 1]$	
	ср. зн. Δ	max Δ	ср. зн. Δ	max Δ
10	30,9	55,8	53,3	65,9
11	27,9	53,2	45,5	66,9
12	29,6	55,8	47,3	67,8
13	27,9	58,0	48,8	68,6
14	29,7	59,8	50,1	69,2
15	31,1	61,3	51,3	69,9
16	28,3	59,7	45,8	70,4
17	29,7	61,2	47,0	70,9
18	28,0	62,4	48,0	71,4
19	29,2	63,5	48,9	71,8
20	30,4	64,5	49,8	72,2
21	28,2	63,4	45,5	72,6
22	29,2	64,3	46,4	72,9
23	27,6	65,2	47,1	73,2
24	28,6	65,9	47,8	73,5
25	29,6	66,6	48,5	73,8
26	27,7	65,8	45,0	74,1
27	28,6	66,5	45,7	74,3
28	27,1	67,1	46,3	74,5
29	27,9	67,7	46,9	74,8
30	28,7	68,2	47,4	75,0
31	27,2	67,5	44,5	75,2
32	28,0	68,1	45,0	75,4
33	26,6	68,6	45,5	75,5
34	27,3	69,0	46,0	75,7
35	28,0	69,4	46,5	75,9

численной задачи:

$$IG(\mathcal{P}) \stackrel{\text{dn}}{=} \frac{w(\text{Opt})}{w(\text{Opt}_{LP})}.$$

Для задач с большим значением разрыва целочисленности оценка (5) может быть слишком грубой.

Нетрудно обнаружить такую задачу \mathcal{P} с относительно малым $IG(\mathcal{P})$, что $R(\mathcal{P}) < G(\mathcal{P})$, как не трудно и привести примеры задач \mathcal{P} , для которых оценка $G(\mathcal{P})$ оказывается существенно более точной, чем оценка $R(\mathcal{P})$. Рассмотрим следующий пример из [7]. Пусть $m = 2^k - 1$, $k \in \mathbb{Z}_+$, $U = \{e_1, \dots, e_m\}$, где элементы e_i множества U — это k -мерные векторы над \mathbb{GF}_2 . Каждый из таких векторов $e_i = (e_{i1}, \dots, e_{ik})$ является двоичным представлением числа i , то есть $i = \sum_{j=1}^k 2^{j-1} e_{ij}$. Пусть $S_i = \{e_j : e_i \cdot e_j = 1\}$, где $e_i \cdot e_j$ обозначает скалярное произведение векторов, то есть $S = \{S_1, \dots, S_m\}$. $w(S_i) = 1$ для $i = \overline{1, m}$. Для такой задачи \mathcal{P} мы имеем следующее дробное покрытие Cvr_{LP} :

$$Cvr_{LP} = \left(\frac{2}{m+1}, \dots, \frac{2}{m+1} \right),$$

$Cvr_{LP} \in \mathbb{R}^m$. $w(Cvr_{LP}) = 2m/(m+1) < w(Opt)$, $w(Gr) = k$. $IG(\mathcal{P}) > 0,5 \log_2 m$.

В табл. 4 приведены нижние оценки на значения $IG(\mathcal{P})$ и $R(\mathcal{P})$, а также значения $G(\mathcal{P})$ для $k = \overline{5, 10}$ (в скобках после значения k указана размерность m задачи). Они показывают, что значение $G(\mathcal{P})$ практически не меняется с ростом k , тогда как $R(\mathcal{P})$ больше, чем $G(\mathcal{P})$, и разница между ними растёт с ростом k .

Таблица 4. $IG(\mathcal{P})$, $R(\mathcal{P})$ и $G(\mathcal{P})$ для задач из [7]

k (m)	5 (31)	6 (63)	7 (127)	8 (255)	9 (511)	10 (1023)
$IG(\mathcal{P}) >$	2,48	2,99	3,49	4,00	4,50	5,00
$R(\mathcal{P}) >$	2,58	3,05	3,53	4,02	4,51	5,01
$G(\mathcal{P})$	1,29	1,30	1,30	1,30	1,30	1,30

Вычислительная сложность оценивания $w(Gr)/w(Opt)$. Заметим, что несмотря на полиномиальную вычислительную сложность нахождения $w(Opt_{LP})$, численное решение задачи линейного программирования, получаемой ЛП-релаксацией ЗПМ, для данных из практических задач уже при $m \geq 4500$ требует слишком много времени [8]. Соответственно, решая ЗПМ значительной размерности, вычислительно трудно найти $R(\mathcal{P})$. Это влечёт и значительную вычислительную сложность алгоритма, реализующего метод ветвей и границ и использующего оценку $R(\mathcal{P})$ при оценивании веса оптимального решения подзадач. Тогда как вычислительная сложность получения $G(s)$ составляет $O(m^2n)$.

Выводы

В работе оценивается точность жадного алгоритма решения индивидуальной задачи о покрытии. Показывается, как в этом случае близость веса решения, получаемого жадным алгоритмом, к весу оптимального решения может быть оценена более точно, чем предполагается его оценкой гарантированной точности. Производится сравнение полученной оценки с другими оценками такого типа. Даются примеры задач, для которых представленная оценка оказывается лучше оценок, полученных с помощью ЛП-релаксации задачи о покрытии множества.

ЛИТЕРАТУРА

1. Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. М. : Мир, 1982. 416 с.
2. Dinur I., Steurer D. Analytical approach to parallel repetition // STOC '14: Proceedings of the forty-sixth annual ACM symposium on Theory of computing, ACM. 2013. P. 624–633.
3. Raz R., Safra S. A sub-constant error-probability low-degree test, and sub-constant error-probability PCP characterization of NP // Proceeding STOC '97: Proceedings of the twenty-ninth annual ACM symposium on Theory of computing. 1997. P. 475–484.
4. Feige U. A threshold of $\ln n$ for approximating set cover // J. ACM. 1998. V. 45, No. 4. P. 634–652.
5. Chvatal V. A greedy heuristic for the set-covering problem // Mathematics of operation research. 1979. V. 4, No. 3. P. 233–235.
6. Slavik P. A tight analysis of the greedy algorithm for set cover // Proceedings of the twenty-eighth annual ACM symposium on Theory of computing. 1996. P. 435–441.
7. Vazirani V. Approximation Algorithms. Springer-Verlag. 2003. 380 p.
8. Umetani S., Yagiura M. Relaxation heuristics for the set covering problem // Journal of the Operation Research Society of Japan. 2007. V. 50, No. 4. P. 350–375.

AN ESTIMATION OF THE GREEDY ALGORITHM'S ACCURACY FOR A SET COVER PROBLEM INSTANCE

A.V. Prolubnikov

Ph.D. (Phys.-Math.), Associate Professor, e-mail: a.v.prolubnikov@mail.ru

Dostoevsky Omsk State University, Omsk, Russia

Abstract. Considering the set cover problem, by modifying the approach that gives a logarithmic approximation guarantee for the greedy algorithm, we obtain an estimation of the greedy algorithm's accuracy for a particular input. We compare the presented estimation to another estimations of this type. We give such examples of the set cover problem instances that the presented estimation significantly improves over linear programming relaxation based estimation.

Keywords: set cover problem, greedy algorithm.

REFERENCES

1. Geri M. and Dzhonson D. Vychislitel'nye mashiny i trudnoreshaemye zadachi. Moscow, Mir Publ., 1982, 416 p. (in Russian)
2. Dinur I. and Steurer D. Analytical approach to parallel repetition. STOC '14: Proceedings of the forty-sixth annual ACM symposium on Theory of computing, ACM, 2013, pp. 624–633.
3. Raz R. and Safra S. A sub-constant error-probability low-degree test, and sub-constant error-probability PCP characterization of NP. Proceeding STOC '97: Proceedings of the twenty-ninth annual ACM symposium on Theory of computing, 1997, pp. 475–484.
4. Feige U. A threshold of $\ln n$ for approximating set cover. J. ACM, 1998, vol. 45, no. 4, pp. 634–652.
5. Chvatal V. A greedy heuristic for the set-covering problem. Mathematics of operation research, 1979, vol. 4, no. 3, pp. 233–235.
6. Slavik P. A tight analysis of the greedy algorithm for set cover. Proceedings of the twenty-eighth annual ACM symposium on Theory of computing, 1996, pp. 435–441.
7. Vazirani V. Approximation Algorithms. Springer-Verlag. 2003. 380 p.
8. Umetani S. and Yagiura M. Relaxation heuristics for the set covering problem. Journal of the Operation Research Society of Japan, 2007, vol. 50, no. 4, pp. 350–375.

Дата поступления в редакцию: 19.10.2019